



May 10, 2000

DSSD CENSUS 2000 PROCEDURES AND OPERATIONS MEMORANDUM SERIES # C - 3

MEMORANDUM FOR John H. Thompson
Associate Director of Decennial Census

Through: Michael J. Longini
Chief, Decennial Systems and Contracts Management Office

From: Howard Hogan *Howard Hogan*
Chief, Decennial Statistical Studies Division

Prepared by: Nick Alberti *na*
Data Processing Techniques and Requirements Staff
Decennial Statistical Studies Division

Subject: Observation of Data Capture Operation in the Baltimore Data
Capture Center on March 30, 2000

I visited the Baltimore Data Capture Center (DCC) on March 30, 2000 to observe data capture operations. Everyone I met at the DCC was very helpful and knowledgeable of the operations. I observed the Audit and Resolution process, the Key From Image (KFI) process, the Key From Paper (KFP) process and the Check-out process.

The Audit and Resolution (AR) staff is very proficient with this review and updating process. They complete cases very quickly. Everyone I observed reviewed all of the person panels on each form and the entire roster when appropriate. No one took shortcuts by skipping those portions of the questionnaire one would expect to be blank. I observed no one applying the AR procedures incorrectly. On the day I observed there was not enough work to keep the AR staff continuously busy and the staff experienced repeated breaks between incoming work. Many cases I observed had stray marks that caused the DCS2000 to code as valid initially, person panels that were not filled out or were crossed out by the respondent. The AR clerks were correctly coding these person panels as invalid.

The KFP unit had plenty of work. Each questionnaire passes through two stages of the KFP operation; an initial keying stage and a quality check stage in which the forms are keyed a second time independent of the initial stage. In the quality check stage the computer compares the entry for each field as it is keyed, with the entry keyed in the initial stage. If the two entries disagree, an error message is shown on the data entry screen and the data entry clerk is asked to key the data two more times.

The KFP software provides a function key that allows the data entry staff to skip over whole pages that are blank but not over blank sections within a page. Within each non-blank page of a form the data entry clerk is required to place an entry (either blank or non-blank) in every response field. The KFP staff dealt with this limitation by holding down the 'Enter' key to quickly place blanks in the blank response fields and proceed to the next field. I was told that the function that allows a data entry clerk to skip over blank pages had been causing errors in the quality check stage of KFP and was not being used by data entry clerks in the second stage. After observing several KFP staff I concluded that the keying software is cumbersome.

The KFI software displays an image of a small portion of the questionnaire page with a field highlighted for keying. During my observation of the KFI, I saw many blank fields highlighted for keying. I observed this happening most often for the questions on race and hispanic origin.

The Check-out operation has two stages. In the first stage an accounting of the data capture status is carried out by scanning the bar code of each form in a batch. Each form is either accepted or rejected based on whether or not data for the form has been successfully captured. The rejected forms are placed on top of the batch and the batch is passed on to a check-out analyst. As the check-out analyst scans the bar code of each rejected form, a message describing the reason for rejection and the suggested disposition of the form is displayed on a monitor. The four possible dispositions of forms are KFP, scanning, manual check-in and supervisor attention. Most forms are sent to KFP or scanning. The check-out analyst may decide to override the suggested disposition of the forms based on the condition of the forms. For example, forms filled out in red ink usually are rejected and the suggested disposition is scanning. The analyst will send these to KFP because they know that scanning will not be successful.

I was told that the most common reasons for rejected forms were image width failure, Optical Mark Recognition/Optical Character Recognition failure, and blank form. I observed the analysis of about twenty short form batches and one long form batch. Most of the rejected forms in these batches met the criteria for a blank form and were sent to KFP. Many of the rejected forms were filled out legibly in pencil or dark ink. These were sent to scanning as suggested by the check-out software. A few forms were filled out in red ink. The suggested disposition of these forms was scanning but the analyst appropriately sent these to KFP.

cc:

DSSD Census 2000 Procedures and Operations Memorandum Series List